Chapter 10

MULTIGRID STRATEGIES

10.1 Motivation

The idea of systematically using sets of coarser grids to accelerate the convergence of iterative schemes that arise from the numerical solution to partial differential equations was made popular by the work of Brandt. There are many variations of the process, which is by no means unique, and many viewpoints of the underlying theory. The viewpoint presented here is a natural extension of the concepts discussed above.

10.1.1 Eigenvector and Eigenvalue Identification with Space Frequencies

Consider the eigensystem of the model matrix $B(\frac{1}{2}, -1, \frac{1}{2})$. The eigenvalues and eigenvectors for $M = 5$ are given in Eqs. 9.53 and 9.52, respectively. Notice that as the magnitudes of the eigenvalues increase, the space-frequency (number of sign changes) of the corresponding eigenvectors also increase. This has a rational explanation from the origin of the banded matrix. Note that

$$\frac{\partial^2}{\partial x^2} \sin(mx) = -m^2 \sin(mx) \tag{10.1}$$

and recall that

$$\delta_{xx}\vec{\phi} = \frac{1}{\Delta x^2} B(1, -2, 1)\vec{\phi} = X \left[\frac{1}{\Delta x^2} D(\vec{\lambda})\right] X^{-1}\vec{\phi} \tag{10.2}$$

We have seen that $X^{-1}\vec{\phi}$ represents a sine transform, and $X\vec{\phi}$, a sine synthesis. Therefore, the operation $\frac{1}{\Delta x^2} D(\vec{\lambda})$ represents the numerical approximation of the multiplication of the appropriate sine wave by the negative square of its frequency, $-m^2$. One finds that

$$\frac{1}{\Delta x^2}\lambda_m = \left(\frac{M+1}{\pi}\right)^2\left[-2+2\cos\left(\frac{m\pi}{M+1}\right)\right] \approx -m^2, m << M \tag{10.3}$$

Hence, the correlation of large magnitudes of λ_m with high space-frequencies is to be expected for these particular matrix operators. However, this correlation is not necessary in general. In fact, the complete counterexample of the above association is contained in the eigensystem for $B(\frac{1}{2}, 1, \frac{1}{2})$. For this matrix one finds, from Appendix , exactly the opposite behavior.

10.1.2 Properties of the Iterative Process

First of all we assume that the difference equations representing the basic partial differential equations are in a form that can be related to a matrix which has certain basic properties. This form can be arrived at "naturally" by simply replacing the derivatives in the PDE with difference schemes, as in the example given by Eq. 3.22, or it can be "contrived" by further conditioning, as in the examples given by Eq. 9.7. These basic properties are:

1. The eigenvalues, λ_m, of the matrix are all real and negative.
2. The λ_m are fairly evenly distributed between their maximum and minimum values.
3. The eigenvectors associated with the eigenvalues having largest magnitudes can be correlated with high frequencies on the differencing mesh.

These conditions are sufficient to ensure the validity of the process described next.

Having preconditioned (if necessary) the basic finite differencing scheme by a procedure equivalent to the multiplication by a matrix C, we are led to the starting formulation

$$C[A_b\vec{\phi}_\infty - \vec{f}_b] = 0 \tag{10.4}$$

where the matrix formed by the product CA_b has the three properties given above. In Eq. 10.4, the vector $\vec{f}_b$ represents the boundary conditions and the forcing function, if any, and $\vec{\phi}_\infty$ is a vector representing the desired exact solution. We start with some initial guess for $\vec{\phi}_\infty$ and proceed through n iterations making use of some iterative process that greatly reduces the amplitudes of the eigenvectors associated with the eigenvalues in the range between $|\lambda|_{max}$ and $\frac{1}{2}|\lambda|_{max}$. We do not attempt to develop

an optimum procedure here, but for clarity we suppose that the three-step Richardson method illustrated in Fig. 9.7 is used. At the end of the three steps we find $\vec{r}$, the residual, where

$$\vec{r} = C[A_b\vec{\phi} - \vec{f}_b] \tag{10.5}$$

Recall that the $\vec{\phi}$ used to compute $\vec{r}$ is composed of the exact solution $\vec{\phi}_\infty$ and the error $\vec{e}$ in such a way that

$$A\vec{e} - \vec{r} = 0 \tag{10.6}$$

where

$$A \equiv CA_b \tag{10.7}$$

If one could solve Eq. 10.6 for $\vec{e}$ then

$$\vec{\phi}_\infty = \vec{\phi} - \vec{e} \tag{10.8}$$

Now we can write the exact solution for $\vec{e}$ in terms of the eigenvectors of A, and the σ eigenvalues of the Richardson process in the form:

$$\vec{e} = \sum_{m=1}^{M/2} c_m\vec{x}_m \prod_{n=1}^{3}[\sigma(\lambda_m h_m)] + \underbrace{\sum_{m=M/2+1}^{M} c_m\vec{x}_m \prod_{n=1}^{3}[\sigma(\lambda_m h_m)]}_{\text{very low amplitude}} \tag{10.9}$$

Combining the properties of the Richardson algorithm and our three basic properties, we can be sure that the high frequency content of $\vec{e}$ has been greatly reduced (about 1% or less of its original value in the initial guess).

Next we construct a permutation matrix which separates a vector into two parts, one containing the odd entries, and the other the even entries of the original matrix (or any other appropriate sorting which is consistent with the interpolation approximation to be discussed below). For example

$$\begin{bmatrix} e_2 \\ e_4 \\ e_6 \\ e_1 \\ e_3 \\ e_5 \\ e_7 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix} \quad ; \begin{bmatrix} \vec{e}_e \\ \vec{e}_0 \end{bmatrix} = P\vec{e} \tag{10.10}$$

Multiply Eq. 10.6 from the left by P and, since a permutation matrix has an inverse which is its transpose, we can write

$$PA[P^{-1}P]\vec{e} = P\vec{r} \tag{10.11}$$

The operation PAP^{-1} partitions the A matrix to form

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} \vec{e}_e \\ \vec{e}_0 \end{bmatrix} = \begin{bmatrix} \vec{r}_e \\ \vec{r}_0 \end{bmatrix} \tag{10.12}$$

Notice that

$$A_1\vec{e}_e + A_2\vec{e}_0 = \vec{r}_e \tag{10.13}$$

is an exact expression. At this point we make our one crucial assumption. It is that there is some connection between $\vec{e}_e$ and $\vec{e}_o$ brought about by the smoothing property of the Richardson relaxation procedure. Since the top half of the frequency spectrum has been removed, it is reasonable to suppose that the odd points are the average of the even points. For example

$$\begin{aligned} e_1 &\approx \frac{1}{2}(e_a + e_2) \\ e_3 &\approx \frac{1}{2}(e_2 + e_4) \\ e_5 &\approx \frac{1}{2}(e_4 + e_6) \qquad \text{or } \vec{e}_0 = A_2'\vec{e}_e \\ e_7 &\approx \frac{1}{2}(e_6 + e_b) \end{aligned} \tag{10.14}$$

It is important to notice that e_a and e_b represent errors on the boundaries where the error is zero if the boundary conditions are given. It is also important to notice that we are dealing with the relation between $\vec{e}$ and $\vec{r}$ so the original boundary conditions and forcing function (which are contained in $\vec{f}$ in the basic formulation) no longer appear in the problem. Hence, no aliasing of these functions can occur in subsequent steps. Finally, notice that, in this formulation, the averaging of $\vec{e}$ is our only approximation, no operations on $\vec{r}$ are required or justified.

If the boundary conditions are Dirichlet, e_a and e_b are zero, and one can write for the example case

$$A_2' = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \tag{10.15}$$

With this approximation Eq. 10.11 reduces to

$$\begin{aligned} & A_c\vec{e}_e - \vec{r}_e = 0 \\ \text{where} \quad & A_c = [A_1 + A_2 A_2'] \end{aligned} \tag{10.16}$$

The form of A_c, the matrix on the coarse mesh, is completely determined by the choice of the permutation matrix and the interpolation approximation. If the original A had been $B(1,-2,1)$, our 7-point example would produce

$$PAP^{-1} = \begin{bmatrix} -2 & & & 1 & 1 & & \\ & -2 & & & 1 & 1 & \\ & & -2 & & & 1 & 1 \\ 1 & 1 & & -2 & & & \\ 1 & 1 & & & -2 & & \\ & 1 & 1 & & & -2 & \\ & & 1 & & & & -2 \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \tag{10.17}$$

and Eq. 10.16 gives

$$\overbrace{\begin{bmatrix} -2 & & \\ & -2 & \\ & & -2 \end{bmatrix}}^{A_1} + \overbrace{\begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix}}^{A_2} \cdot \frac{1}{2} \overbrace{\begin{bmatrix} 1 & \\ 1 & 1 \\ & 1 \\ & & 1 \end{bmatrix}}^{A_2'} = \overbrace{\begin{bmatrix} -1 & 1/2 & \\ 1/2 & -1 & 1/2 \\ & 1/2 & -1 \end{bmatrix}}^{A_C} \tag{10.18}$$

This process is deceptively simple. We started with the equation $B(1,-2,1)\vec{e} = \vec{r}$ on the fine mesh and reduced the problem to the equation $\frac{1}{2}B(1,-2,1)\vec{e}_e = \vec{r}_e$ on the next coarser mesh. It appears as if the data on the odd points had been ignored altogether and a scaling factor had arbitrarily appeared. Such is not the case, however, and except for the assumption in Eq. 10.14 the process is quite rigorous.

If the boundary conditions are mixed Dirichlet-Neumann, A in the 1-D model equation is $B(1, \vec{b}, 1)$ where $\vec{b} = [-2, -2, ..., -2, -1]^T$. The eigensystem is given by Eq. B.22. It is easy to show that the high space-frequencies still correspond to the eigenvalues with high magnitudes, and, in fact, all of the properties given in Section 10.1 are met. However, the eigenvector structure is different from that shown in Fig. 9.4 for Dirichlet conditions. In the present case they are given by

$$x_{jm} = \sin\left[j\left(\frac{(2m-1)\pi}{2M+1}\right)\right] \quad ; \quad m = 1, 2, \cdots, M \tag{10.19}$$

and are illustrated in Fig. reffig:fig10.1. All of them go through zero on the left (Dirichlet) side, and all of them reflect on the right (Neumann) side, being symmetrical about the point $m = +\frac{1}{2}$ where $x = \pi$ and their magnitude is 1.

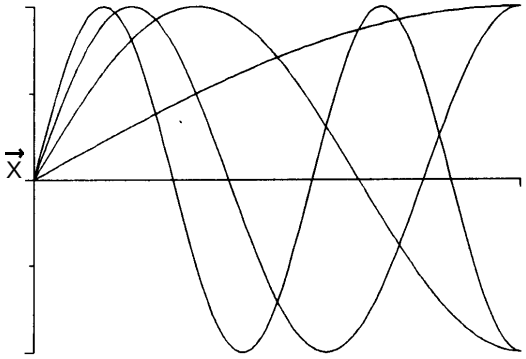

Figure 10.1: Ėigenvectors for the mixed Dirichlet–Neumann case.

For Neumann conditions, the interpolation formula in Eq. 10.14 must be changed. In the particular case illustrated in Fig. 10.1, e_b is equal to e_M. If Neumann conditions are on the left, $e_a = e_1$. When $e_b = e_M$, the example in Eq. 10.15 changes to

$$A_2' = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \tag{10.20}$$

The permutation matrix remains the same and both A_1 and A_2 in the partitioned matrix PAP^{-1} are unchanged (only A_4 is modified by putting -1 in the lower right element). Therefore, we can construct the coarse matrix from

$$\overbrace{\begin{bmatrix} -2 & & \\ & -2 & \\ & & -2 \end{bmatrix}}^{A_1} + \overbrace{\begin{bmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & 1 \end{bmatrix}}^{A_2} \cdot \frac{1}{2} \overbrace{\begin{bmatrix} 1 & & \\ 1 & 1 & \\ & 1 & 1 \\ & & 2 \end{bmatrix}}^{A_2'} = \overbrace{\begin{bmatrix} -1 & 1/2 & \\ 1/2 & -1 & 1/2 \\ & 1/2 & -1/2 \end{bmatrix}}^{A_C} \quad (10.21)$$

which gives us what we might have "expected" and shows us that the process is recursive.

The remaining steps required to complete an entire multigrid process are relatively straightforward, but they vary depending on the problem and the user. The reduction can be, and usually is, carried to even coarser grids before returning to the finest level. However, in each case the appropriate permutation matrix and the interpolation approximation define both the down- and up-going paths. The details of finding optimum technique are, obviously, quite important but they are not discussed here.